Lecture 5

02/01/2022

- ## Twin pillars of statistical inference

Classical / Frequentist

Probabilities refer to the frequency of outcomes in repeated experiments

Bayesian

Probabilities refer to subjective belief conditioned on data and prior knowledge.

$\Rightarrow$ PRIORS in Bayesian statistics act as boundaries on our inference.

- ## Maximum Likelihood Estimation

1. Formulate a model with some parameters, θ.
2. Find the parameters θ^* that best fit the observed data.
3. Estimate the uncertainty on those parameters.
4. Try other models, and see which fits best.

- <u>Likelihood</u> $\longrightarrow$ fitting/objective function.

e.g. let's assume the noise in our data is Gaussian.

$$\text{data} = \text{signal} + \text{noise}$$
$$d = s + n$$

Gaussian

$$p(n_i) = N(0, \sigma_i)$$

$$\therefore \quad p(d_i) = N(s_i, \sigma_i)$$

$$\underbrace{\text{LIKELIHOOD}_i}_{\text{PROBABILITY OF DATA GIVEN PARAMETERS}} = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2}\frac{(d_i - s_i)^2}{\sigma_i^2}}$$

$$\Rightarrow \begin{pmatrix}\text{LIKELIHOOD OF} \\ \text{ALL DATA}\end{pmatrix} = \prod p(d_i)$$
$$= \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2}\sum_i \frac{(d_i - s_i)^2}{\sigma_i^2}}$$

$\boxed{\text{NOTE}}$ $$\chi^2 = \sum_i \frac{(d_i - s_i)^2}{\sigma_i^2}$$

... MAXIMIZING the (log)-likelihood is the SAME as MINIMIZING the χ^2 value.

MLE

$$\boxed{\text{MLE}} \xrightarrow{\text{HOMOSCEDASTIC}} \quad \hat{\mu} = \frac{1}{N}\sum_i^N \overbrace{x_i}^{\text{data}}, \qquad \sigma_{\hat{\mu}} = \frac{\sigma}{\sqrt{N}}$$

$$\boxed{\text{MLE}} \xrightarrow{\text{HETEROSCEDASTIC}} \quad \hat{\mu} = \frac{\sum_i^N \left(x_i / \sigma_i^2\right)}{\sum_i^N \left(1/\sigma_i^2\right)}, \qquad \sigma_{\hat{\mu}} = \left[\sum_i^N \left(1/\sigma_i^2\right)\right]^{-1/2}$$

Quantifying Uncertainty on Model Parameters

* With >1 parameter, you can have correlated measurement uncertainties.

$$\sigma_{jk} = \left(\left[F^{-1}\right]_{jk}\right)^{1/2}$$

where

$$\underbrace{F_{jk} = -\left.\frac{\partial^2 \ln L}{\partial\theta_j \partial\theta_k}\right|_{\theta = \theta_{MLE}}}_{\text{FISHER INFORMATION MATRIX}}$$

e.g. Measuring the mean μ of Gaussian-distributed data with homoscedastic uncertainties ... $\sigma_{\hat{\mu}} = \frac{\sigma}{\sqrt{N}}$